

seminar series240

How does PISA shape
education policy making?
Why how we measure
learning determines what
counts in education

Simon Breakspear

Centre for Strategic Education

(CSE) is the business name for IARTV
ABN 33 004 055 556

Mercer House 82 Jolimont Street
East Melbourne Victoria 3002

Phone +61 3 9654 1200

Fax +61 3 9650 5396

Email info@cse.edu.au

www.cse.edu.au

240

How does PISA shape education policy making? Why how we measure learning determines what counts in education

Simon Breakspear

Introduction	3
Section One	
What PISA measures: A global yardstick for learning	4
Section Two	
How PISA shapes policy: Three emerging tensions	6
Section Three	
Implications for system reform: Putting PISA in its place	11
Conclusion: Preferable PISA futures	13

ISSN 1838-8558
ISBN 978-1-921823-62-6

© 2014 Centre for Strategic Education, Victoria.

The Centre for Strategic Education* welcomes usage of this publication within the restraints imposed by the Copyright Act. Where the material is to be sold for profit then written authority must be obtained first. Detailed requests for usage not specifically permitted by the Copyright Act should be submitted in writing to:

The Centre for Strategic Education
Mercer House, 82 Jolimont Street,
East Melbourne VIC 3002.

(*The Centre for Strategic Education (CSE) is the business name adopted in 2006 for the Incorporated Association of Registered Teachers of Victoria (IARTV). Therefore, publications which were previously published in the name of IARTV are now published in the name of CSE.)

Produced in Australia by Centre for Strategic Education
Mercer House, 82 Jolimont Street, East Melbourne VIC 3002

Editorial Team: Tony Mackay, Keith Redman,
Murray Cropley, Andrew Miller

Introduction

What educational outcomes we measure, and how we measure them, affects our understanding of the performance of our systems and the policy action that is taken for improvement. The vast majority of analysis and debate in the field of whole-system reform is centred on how to improve the **means** of educational change. Too often the **ends** of education – and their measurement – are assumed. Yet, those interested in improving school systems must answer an important question: how should judgements be made about the performance and progress of a school system over time? Or, to put it another way: how can we ensure that the education system is achieving its aims and purposes over time?

Increasingly, system leaders have accepted the Organisation for Economic Cooperation and Development's (OECD) Programme of

International Student Achievement (PISA) as the global yardstick for school system performance and progress over time. First initiated in the year 2000 (see OECD, 2000), PISA has gained an almost 'taken-for-granted' status in public and policy debates about education quality. PISA is a triennial international comparative study of student learning outcomes in reading, mathematics and science. With over 60 systems regularly participating in the survey, the release of the PISA international performance league table attracts substantial attention from both the media and policy makers, and at times has even triggered large-scale reform. PISA is increasingly accepted as a proxy for school system performance, and PISA high-performing countries are now regularly referred to in policy circles as simply high-performing school systems.

In this paper I focus on the policy impacts of PISA. I argue that PISA has become the gold-standard instrument for evaluating the performance of education systems globally. As a consequence, PISA has the capacity to shape and frame how national policy makers define and measure the end-goals of their education systems, judge system performance and enact policy reform. The end-goals of education systems, such as improved quality and equity, are multidimensional, ambiguous and often highly contested. I seek to draw attention to the fact that PISA indicators function to simplify and quantify complex educational phenomena and thereby set new standard definitions for how educational goals are understood, communicated and tracked, with important normative implications: what we

choose to measure in education shapes what we collectively strive to achieve.

The paper proceeds in three sections. First, I provide an overview of the PISA survey, its inception, design and growth. In Section 2, I examine the impact of PISA on policy-making processes and highlight three emerging tensions that must be navigated. In Section 3, I explore three key implications for system leaders to ensure that PISA is appropriately positioned in efforts to evaluate and improve school systems. It is worth noting from the outset that this paper does not engage with the scholarly critiques of the reliability or validity of PISA as an assessment instrument. These are important debates,¹ but lie outside the scope of this piece, which focuses on PISA's policy uses and effects.

Section One

What PISA measures: A global yardstick for learning

The emergence of PISA: Filling a gap in international indicators

The development of PISA at the OECD is best understood within its broader agenda of work, establishing frameworks and indicators for comparing education systems internationally during the late 1980s and early 1990s. In 1988 the governing body of the OECD Centre for Educational Research and Innovation (CERI) established the Indicators of National Education Systems (INES) project, which sought to develop and publish cross-national indicators on education (Bottani, 1996; Henry et al, 2001). The OECD's annual *Education at a Glance* publication reported the comparative indicators across member countries. The majority of the comparative indicators represented system level inputs or processes, such as financial resources, access rates and school organisation data. Some output indicators could be generated,

such as student progression rates, but these were not thought of as informative enough to indicate student achievement outcomes. The first editions of *Education at a Glance* (OECD, 1992; 1993) did include some student-learning outcome data, but these were drawn from already existing international assessments that were administered by International Association for the Evaluation of Educational Achievement (IEA) and later the Educational Testing Service (ETS), which had initiated the International Assessment of Educational Progress in 1988 (Bottani, 1996).

The idea of PISA stemmed from what its proponents claimed was a lack of quality and coverage of cross-national student achievement data generated by the IEA and the International Assessment of Educational Progress studies (Bottani, 1996). It became increasingly clear to both OECD Secretariat staff and country representatives on the OECD education committee, for example, that the data lacked quality, were not regularly collected and were

not available for all member countries (Martens and Wolfe, 2004; McGaw, 2008a; Jakobi and Martens, 2010). A strategy for a large-scale assessment was developed, which led eventually to the PISA framework. PISA would fill the perceived gap in the comparative indicator set by providing a measure of system outcomes in terms of student learning.

The first released report of PISA results, *Knowledge and Skills for Life* (OECD, 2001), outlined the project's broad goals:

PISA aims at providing a new basis for policy dialogue and for collaboration in defining and operationalising educational goals – in innovative ways that reflect judgments about the skills that are relevant to adult life. It provides inputs for standard setting and evaluation; insights into the factors that contribute to the development of competencies and into how these factors operate in different countries, and it should lead to a better understanding of the causes and consequences of observed skill shortages. By supporting a shift in policy focus from educational inputs to learning outcomes, PISA can assist countries in seeking to bring about improvements in schooling and better preparation for young people as they enter an adult life of rapid change and deepening global interdependence.

(OECD, 2001, p 3, emphasis added)

PISA was a step change from the OECD's previous educational indicator work, which mostly involved aggregating national data about educational inputs. In contrast, PISA aimed to evaluate the performance of systems and thus required the OECD-curated experts to make judgements regarding how to measure the end-goals of education systems and how to generate a process for collecting new data across systems. The decisions made by OECD-curated experts during this initial design and production of PISA, engaged with important ethical questions about what and who education is for, what counts as an educated person or society and what can and should be measured.

How PISA defines learning outcomes

The central objective of PISA is to provide cross-nationally comparable evidence of student performance on the skills that are judged to be important for adult life, in the context of globalisation and the emerging knowledge society (OECD, 2004; Schleicher, 2007). PISA is designed to assess the extent to which students have the capacity to use the knowledge and skills that they have learnt to meet real-life challenges (OECD 1999). PISA adopted a unique definition of literacy that focuses on the

students' capacity to extrapolate from what they have learned and apply their knowledge in real-life settings, and their capacity to analyse, reason and communicate effectively as they pose, interpret and solve problems in a variety of situations.

(OECD, 2001, p 22)

The central objective of PISA is to provide cross-nationally comparable evidence of student performance on the skills that are judged to be important for adult life

This was a different approach from that taken by the Evaluation of Educational Achievement (IEA) studies – Progress in International Reading Literacy Study (PIRLS) and Trends in International Maths and Science Study (TIMSS)² – which are designed to test the internationally common elements across the curricula of participating countries. The target population for PISA is 15-year-olds who are currently enrolled in school.

PISA specifically assesses student achievement in reading, mathematics and science literacies. For each survey, one of the three testing areas is selected as the major domain; the other two areas are the minor domains and have fewer items in the survey. In the first three-surveys in the PISA 2000–2006 cycle, reading was the major domain in 2000, mathematics in 2003 and science in 2006. The same pattern has been used for the second PISA cycle from

2009–2015. In 2009 reading was the major domain, in 2012 mathematics was the major domain, and in 2015 science is planned as the major domain. Each survey round also assesses additional cross-curriculum competencies beyond reading, mathematics and science (Turner and Adams, 2007). For example, problem solving was assessed in PISA 2003 and in PISA 2009 electronic literacy was assessed. Only a subset of countries selects to participate in these additional assessments.

The PISA survey instrument is conducted as a two-hour test. There is a diversity of test item types used in the PISA surveys that range from multiple-choice items to extended-response questions. Students also complete a 30-minute student background questionnaire, with questions that focus on the students' personal socioeconomic backgrounds, learning outside school, and learning habits and attitudes.

Background questionnaires are also filled out at the school level to provide information about the teaching and learning environment, including resourcing. The information is used to analyse the effects of these student-level and school-level variables on learning outcomes.

The growth and development of PISA

While PISA was initially developed to measure the performance of OECD education systems, an increasing number of non-OECD partner countries and economies take part. Overall, 43 countries took part in 2000, 41 in 2003, 57 in 2006, and over 60 in both 2009 and 2012. Countries range from the 34 highly industrialised OECD countries, the large emerging economies of Brazil, Russia and China, to a vast range of middle-income countries, such as Columbia, Indonesia and Kazakhstan.

Section Two

How PISA shapes policy: Three emerging tensions

As the media, policy makers and the public increasingly accept PISA as a robust and legitimate proxy for education system performance, then PISA has the potential to shape the way in which they judge system performance, define policy problems, set improvement targets and enact policy borrowing and reform. In short, PISA indicators become the lens through which we come to understand our systems.

This section outlines three emerging tensions for policy making, arising from the use and potential misuse of PISA indicators. These are

- the politics of reform;
- the definition of policy problems and targets; and
- the direction of policy learning.

The politics: Catalyst for sustainable reform OR rank-seeking policy action?

The highly publicised PISA evaluation of system performance can act as an external shock in the political economy of education reform. Andreas Schleicher, the head of PISA, highlighted the potential pressure that PISA can generate:

In 1995, at the first meeting of the OECD ministers I attended, every country boasted of its own success and its own brilliant reforms. Now international comparisons make it clear who is failing. There is no place to hide.

(Quoted in The Economist, 2008)

The public pressure caused by media attention on PISA can also exacerbate the reform debate. Every three years, the release of the PISA league table prompts a global discussion about school reform in both international (for

example *The Economist*, *New York Times*) and national media across many OECD and partner countries/economies, with such headlines as:

- ‘An international report card: Shanghai’s school students out-perform all others’ (*The Economist*, 2010);
- ‘A league table to worry us all: The decline of the UK’s PISA scores raises questions about the direction of educational policy’ (Mortimore, 2008, in *The Guardian*); and, in Germany,
- ‘Bad luck, the new educational disaster’ (*Der Spiegel*, 2002).

PISA system outcome indicators seem to be accorded a high level of credibility and thus policy influence, across countries. They are often used as evidence to argue that national performance requires improvement overall or in a specific domain area, serving to overcome the inertia of the political economy of reform. However, this external shock may be utilised to rush through ill-conceived reforms that will attempt to increase PISA rank or scores but have little chance of improving teaching and learning in a sustainable way.

In response to lower-than-expected results on the first two PISA surveys in 2001 and 2004, a few countries enacted large-scale policy changes. This phenomenon became known as ‘PISA Shock.’

Germany, Denmark and Japan have been the most often-cited examples of the policy impact of PISA. In Germany, the education policy debate was intense and changes in light of PISA 2000 were substantial. Ertl (2006), for example, has argued that the PISA-inspired debate over public education resulted in a range of significant reform measures, including the generation of national standards and the provision of further support for disadvantaged students, especially those from immigrant backgrounds (Ertl, 2006).

Whilst less publicised, Egelund (2008) has highlighted that a similar reaction to PISA 2000 occurred in Denmark. Takayama

(2008) identifies Japan as a country that also experienced a ‘PISA shock’ in the national politics of education reform. Japan was a top performer in PISA 2000, yet PISA 2003 showed a decline in its performance, which resulted in a perceived ‘crisis’ that prompted significant public and political debate on education reform. In response, the Ministry for Education reversed a contentious *yutori* (low pressure) curriculum policy and enacted changes to national assessment practices (Takayama, 2008).

As a result of PISA’s prominence, and its perceived authority as an evaluation instrument, it seems that below-expectation results have the potential to enact an external shock in national policy agendas. Kingdon’s (1984) theory of policy agendas may help explain the role of PISA in these large-scale reform agendas. Kingdon proposed that an external shock could generate a ‘policy window’ for reform. This so-called window only opens up the potential for a reform agenda to be established. Yet, actual policy reform would only occur if and as national actors use the opportunity afforded by lower-than-expected PISA results to legitimise the need for large-scale reform.

The publication and focus by the media on average country PISA performance and relative rank on the league-table is creating a policy environment that is not necessarily conducive to evidence-informed policy making. PISA has become a high-stake assessment for policy makers, and has the potential to trigger rank-seeking reforms, with potential negative effects on the quality of teaching and learning. Campbell’s law suggests that ‘the more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures, and the more apt it will be to distort and corrupt the social processes it is intended to monitor’ (Campbell, 1976).

As PISA’s ‘country report card’ is made public via the league table performance every three years, policy leaders are under growing pressure to sustain and improve PISA trend performance and rank. Whilst only a few countries across

the OECD have made measurable progress across domains, there are many that have registered substantial declines. This may result in increasing pressure in these systems to focus rapid reforms on improving PISA rank or score. Yet, this is unlikely to result in coherent and sustainable improvements in teaching and learning.

**Defining policy problems and targets:
Creating clarity OR limiting the end-goals of education?**

PISA defining policy problems

The most important effect of PISA indicators is that they frame issues and define policy problems. Educational outcomes are often ambiguous and difficult to track over time. PISA indicators simplify and quantify the policy problems to be solved in national education systems, a process which involves simplifying complex educational outcomes into aggregated metrics that are easy to understand and use. PISA tracks trends over time, reporting every three years whether a system's performance across reading, mathematics and science is improving, declining or staying stable both in objective terms and relative to other countries. By doing this, PISA is shaping and framing how policy makers understand the performance of their system, and the areas that require focus.

constant high-stakes comparison, along the same narrow indicators, has the potential to limit the national view of what matters educationally

PISA performance outcomes vary substantially across participating countries. The top five performing countries and economies in the PISA 2009 reading assessment were Shanghai–China (556 score points), Korea (539), Finland (536), Hong Kong–China (533) and Singapore (526). The lowest performing countries were Kyrgyzstan (314), Azerbaijan (362), Peru (370), Panama (371) and Qatar (372) (OECD, 2010). As a guide to what the scores mean, the OECD refers to 40 PISA points as equivalent to one year's worth of education.

Indicators define how we understand educational goals. Over time, the indicators become equated with the initial goals that they sought to measure. So, improving educational quality becomes synonymous with improving PISA scores. Or, decreasing the PISA inequity indicator score equates with improving system educational equity. Yet, in simplifying and quantifying complex phenomena and multifaceted constructs by equating an indicator with the original aim, many elements may be lost. For example, the OECD equity indicators of the variance between schools, or the variance in scores accounted for by student social economic background, defines the problem of educational equity to be solved. Making progress on educational equity comes to mean improving these PISA equity indicators. Yet, these indicators do not fully represent the complex social phenomenon of inequity or the broader theoretical construct of equity.

Overall, constant high-stakes comparison, along the same narrow indicators, has the potential to limit the national view of what matters educationally.

PISA-based system improvement targets

One consequence of this shift is that PISA is increasingly being used to evaluate reforms and set future performance benchmarks for improvement. My own research across PISA-participating countries has identified three common types of PISA-based targets across countries (Breakspear, 2012). These are

1. the relative rank of a country in international performance league tables (for example, to achieve in the top five countries). These types of targets are dependent on the relative performance of other countries.
2. specific national PISA scores, such as a numeric performance score that should be reached, often set as the OECD average of 500 PISA points; and
3. equity goals, including decreases in the percentage of students scoring at or below PISA proficiency Level 2 (the PISA baseline proficiency level) or a reduction in the variance between schools.

There are growing numbers of prominent examples of countries setting explicit rank-focused targets for PISA improvement. Australia has set the goal to be ranked in the top five countries in the PISA league tables by 2025. This new target may have been a response to Australia's significant decline in PISA reading performance between PISA 2000 and PISA 2009. In January 2010 average-performing Denmark set a performance target of achieving in the top five in PISA. In PISA 2009, Denmark ranked 24th in reading. So too, the Minister for Education in Wales set a target for Wales to be in the top 20 leading nations for PISA 2015. In PISA 2009 Wales ranked 38th for reading, 40th for mathematics and 30th for science.

Stanley has questioned the utility of setting rank-based targets.

Setting national or local education targets based on rank position on international tests, rather than on specific standards achieved, flies in the face of modern assessment practice, which is standards-based. Rank position is inappropriate as a goal for improvement, as national rankings can be influenced by quite small differences in student cohort scores. Such differences may not, in themselves, represent meaningful learning and/or skill differences in standards attained, which should be the main focus. Despite these concerns, rank position appeals to those who see improving educational outcomes as a competitive sport between nations.

(Stanley, 2013, in his foreword to Masters, 2013, p iii)

Other countries have committed themselves to specific PISA score targets. For example, both Thailand and Brazil have set the PISA performance target of reaching the OECD average performance by 2021, which is always set by the OECD at 500 PISA points. Brazil also implemented a comprehensive national monitoring system where all schools and regions are rated on the Index of Development

of Basic Education (IDEB). This Index was benchmarked to PISA in order to allow the government to track school-level progress regularly towards the 2021 PISA target. Finally, in Mexico, as part of an Education Sector Programme 2007–2012, a range of objectives and performance targets were set, including the goal of raising PISA performance for both mathematics and reading to 435 by 2012.

Finally, other systems have focused on PISA equity targets. For example, the EU included a PISA-based target for low achievers within its strategic framework for European cooperation on education and training ('ET 2020'). The target states that by 2020 '*the share of 15-year-olds with insufficient abilities in reading, mathematics and science should be less than 15%*' (Commission of the European Communities, 2011). This target is explicitly linked to PISA, with 'insufficient abilities' being defined as the share of pupils performing below PISA proficiency Level 2.

There are growing numbers of prominent examples of countries setting explicit rank-focused targets for PISA improvement.

Overall, the embedding of PISA policy improvement targets indicates policy makers' views that an education system's direction of improvement should be judged by PISA evaluation. The emerging logic is that if a reform initiative is to be judged as successful in improving the quality of an education system then it must be evidenced, at least partly, by an improvement in PISA indicators.

Setting PISA targets may support reform efforts by clarifying clear and measurable targets for improvement and thereby mobilising system actors around shared goals. Alternatively, it is possible that such targets only increase the level to which PISA becomes a 'high-stakes' assessment for policy makers, leading to 'rank-seeking' policy actions that are not aligned with coherent and sustainable reform.

Policy learning – New insights OR reduced curiosity

PISA measurement influences policy through directing education policy borrowing and learning. There is a long history of education policy borrowing and lending between countries (Phillips and Ochs, 2004). For example, during the 20th century there was substantial mutual interest and analysis of the education systems of the USA, Japan, Germany and England. Acting as an educational Rosetta Stone, PISA has catalysed a new global dialogue regarding the design and outcomes of different education systems. Having a common yardstick of quality has encouraged policy discussion between countries that are culturally and linguistically distinct, and thereby enabled new learning.

PISA measurement influences policy through directing education policy borrowing and learning.

The identification of high-performing education systems through PISA is shifting the direction of education borrowing. The construction of PISA indicators and the subsequent rating and ranking of countries against common scales leads to new classifications of education systems. PISA classifies systems as high, average or low-performing education systems. Over successive survey rounds, the OECD has identified consistent high PISA performers – Finland, Canada, Japan, and Korea. These countries act as new reference societies and the target of policy borrowing. Schleicher states the following.

Results from PISA show that strong educational performance, and indeed significant improvement over short periods of time, is possible. Whether in Asia (eg, Japan and Korea), in Europe (eg, Finland) or in North America (Canada), many countries display strong overall learning outcomes; equally important, they show that poor performance in school does not automatically follow from a disadvantaged socio-economic background.

Furthermore, some countries show that educational success can become consistent and predictable, with very little variation in performance across schools. These countries set the goals to which others can aspire.

(Schleicher, 2009, p 251)

The classification of education systems by PISA has implications for both normative pressure to reform and policy-borrowing behaviour. Average and low-performing systems are regarded as countries where reform efforts must occur in order to improve performance. These systems become the recipients of policy lessons that should be learned from the consistently high performers, or the successful PISA improvers. For example, the top performance of Finland across the first 4 PISA surveys led to a greater interest in understanding the causes that underlie the success of the Finnish education system and the potential application of these practices for system improvement elsewhere (Sahlberg, 2011; Takayama, 2010). Since PISA 2009, Shanghai has attracted increasing attention as a source of potential policy borrowing (Tucker, 2011)

There is a need for caution in two areas. First, the need for system leaders to utilise PISA to engage in policy learning rather than direct copying. There has been a growing misguided tendency for some policy makers to ‘cherry-pick’ individual policies from high-performing PISA countries with little regard for cultural context or overall policy coherence. Secondly, PISA may have initially stimulated interest in education systems previously overlooked, such as Finland, but it may now also begin to limit the scope of sources of learning and inspiration. PISA is generating a standardised list of countries that are categorised as appropriate policy learning targets. It is worth considering, in the age of PISA, could a policy leader announce that their system was seeking to learn from another country that was rated as a lower PISA performer?

Section Three

Implications for system reform: Putting PISA in its place

In the previous section I have outlined how PISA has the capacity to shape how national policy actors think about the role of education systems in the 21st century. In this section, I outline three key strategies to avoid PISA misleading policy. I argue that PISA should be positioned as a resource for policy-making rather than a high-stakes public evaluation with the potential to cause unintended consequences in reform processes. There are actions for both policy leaders and the OECD itself in order to ensure PISA plays a positive role in coherent and sustainable evidence-informed reform processes.

Implication 1: Clarify educational goals and purposes

The end-goals of education

The potential problem of PISA's prominence as a proxy for overall school system effectiveness is that its narrow indicators can become equated with the education system end-goals themselves. PISA may produce useful metrics of the performance of 15-year-olds on a particular conceptualisation of reading, mathematics and science skills, but it cannot represent the broader goals and purposes of school systems. Simply seeking to improve PISA scores is an anaemic vision for the education of young people. Narrow indicators should not be equated with the end-goals of education.

The discussion of educational end-goals involves ethical deliberation about what matters in education and what an educated person should be. It is important for education systems to consider and state clearly what they value educationally. We must engage in an ethical discussion – what Pring describes as a ‘deliberation about what is worthwhile in the development of persons’ (2012, p 29). These are questions that must be debated in the public sphere through democratic engagement, not

merely technocratic expert decision making. Democracies must ask: Does PISA effectively represent and measure what we want for all of our young people?

Many systems around the world are beginning to make public commitments regarding the educational attainment goals desired for all of their students. In Australia the Melbourne declaration states the goal that all young people become: successful learners, confident and creative individuals, and active and informed citizens (MCEETYA, 2008). Alberta, Canada provides a similar example. Under the Inspiring Education reform, a province-wide consultation was conducted in order to determine the capabilities young Albertans should have by 2030. As a province, Alberta has determined that all students will be given the learning experiences to enable them to become engaged thinkers and ethical citizens, with an entrepreneurial spirit.³ Both these examples are compelling statements about the end-goals of education that extend far beyond the narrow range of competencies that PISA can assess.

The purposes of education

It is also important for system leaders to consider the broader purposes of education and the appropriateness of using PISA to guide policy. All indicators are developed on the basis of particular theoretical models and assumptions. These foundational ideas

Narrow indicators should not be equated with the end-goals of education.

are often not made explicit in the indicator. Poovey (1998) has argued that a distinctive feature of modernity is the use of numerical information to frame and represent the world in ways that appear to be objective and values-free, but that these representations actually obscure the theoretical assumptions and ideas. PISA is steeped in the ideas of human capital

theories of economic growth, and the changing skills needed for national and individual competitiveness. In the mid-1990s the OECD adopted, and to some extent conceptually developed, an economic and human capital lens on the purposes of education, and the concept of lifelong learning. Meanwhile, alternative theoretical approaches were also being espoused by other global education organisations that emphasised a rights-based rather than human capital approach to education. For example, the OECD economic conceptualisation of the role of education and its view of lifelong learning differed significantly from the humanist view advocated by UNESCO, which emphasised learning for social and democratic engagement and participation (Delors, 1996). The UNESCO *Learning the Treasure Within* report (Delors, 1996) highlighted the purposes of learning as learning to know, learning to do, learning to live together and learning to be (Delors, 1996).

It is important to recognise that within PISA indicators, certain education purposes are absent, substantially diminished or reframed around economic ends. For example, the role of education in national cultural and civic socialisation is ignored in PISA indicators.

It is important to recognise that within PISA indicators, certain education purposes are absent, substantially diminished or reframed around economic ends.

Furthermore, educational inclusion and equity, which have deep conceptual roots within the rights-based discourses of education, have been reframed within a human capital perspective, focusing on ensuring all future workers maximise productive output by having their skills fully optimised. Additionally, investment in education is argued for on the basis that a future low-skill adult will be an economic liability. This is not to say that PISA does not have relevance. It is important, however, for system leaders to consider the level of alignment between the theoretical underpinning of PISA

and their own values and assumptions about the purposes of education and the identity of learners. An over-reliance on PISA may lead to decisions being made from a human capital approach to education.

Implication 2: Seek broad metrics beyond PISA

The second key implication for system leaders is that they should seek a broad range of metrics beyond PISA. PISA indicators are simply too narrow to represent the broader range of goals that school system are tasked with developing across economic, social, civic and human development domains. PISA indicators only measure a narrow range of cognitive skills against a particular conceptualisation of learning outcomes. It also only measures a narrow conceptualisation of reading, mathematics and science, due to testing time limitations. The end result is that PISA is a very narrow measure on which to judge the overall performance of an education system, and provides insufficient information for designing improvement strategies.

Education systems are complex and multi-dimensional. No reform should be enacted or evaluated based on a single assessment measure. Furthermore, mismeasuring our education systems could lead to misleading policy making. If PISA becomes the sole metric of focus then policy may be shifted to optimise PISA scores, with detrimental trade-off effects on other important education goals that are not measured.

There is a need for system leaders to move beyond PISA in order to generate a more appropriate set of metrics from which to understand and govern education. First, we must move away from having one gold standard measure – PISA – and rather harness insights from a broader range of currently available data through IEA studies of mathematics, reading and citizenship, along with rich national data. Secondly, there must be an investment in the research and development necessary to generate new indicators across

the broader range of educational goals that a system commits to develop in its children and youth. The PISA 2015 collaborative problem-solving assessment is an important attempt to broaden the testing focus of PISA to include the important area of non-cognitive skills. Yet, this is only the beginning. We must develop the system capacities to measure what matters in educational progress, especially in the intrapersonal skills of resilience, grit and learning agility. Thirdly, we should place renewed emphasis on indicators of student wellbeing, health and engagement. Finally, there should be an acknowledgement that some important aspects of education will remain elusive to empirical measurement, yet must still be emphasised and valued.

Pursue deeper policy analysis

PISA should not be used as a definitive judgement of education system performance. Rather, PISA evaluation should be the starting point for a diagnostic approach to understanding an education system, whereby multiple sources of data are used to develop a more complete picture of the system. Large-scale system assessments should be coupled with detailed qualitative work that explores what is really happening at the level of teaching and learning. It is of course possible that strong outcomes indicators may not always represent high-quality teaching and learning. So too, high-quality teaching may not always show strong outcomes in areas facing local complexities or disadvantage.

PISA has an explicit policy orientation in its design. Beyond measuring and monitoring outcomes, PISA was also intended to provide policy insights into the factors that may account for differential performance within and between countries and thus serve in more effective policy making. Here the OECD should be explicit about the limits of PISA – as a narrow metric devoid of cultural or contextual insights – as a tool for policy action. Policy leaders need to be mindful and aware that measurements, even from authoritative organisations like the OECD, have their limits.

Too often policy makers fail to differentiate between correlation and causation in the policy findings generated from PISA analysis. Furthermore, there is still a tendency to assume policies generated in one context will be successful in another. There is a need to educate explicitly the users of PISA indicators – both policy makers and the media – in order to ensure PISA data is not used for purposes beyond the scope of what the survey can validly claim.

Too often policy makers fail to differentiate between correlation and causation

As an alternative to over-relying on the policy analysis generated from the PISA survey alone, there is a huge opportunity to harness insights from the rich country-specific policy work conducted within other divisions of the OECD Directorate of Education and Skills, such as the comprehensive review of assessment and evaluation policies across more than 20 countries (OECD, 2014). Furthermore, the work of OECD CERI describing rich case studies of innovative learning environments across OECD countries provides powerful insights into new forms of learning design with substantial potential to improve student outcomes (OECD, 2013).

Conclusion: Preferable PISA futures

PISA has sparked unparalleled global discussion between educational researchers, policy makers and practitioners. Yet, the growing dominance of PISA as the key proxy for determining system performance and the high-stakes consequences for policy leaders, have potential negative consequences on policy-making processes. Throughout this paper I have argued that how we choose to measure educational system progress will define and shape what policy-makers and the public come to understand to be educationally valuable.

It is time to put PISA in its place. I am not arguing that PISA should be ignored, nor that PISA should be dismantled. The problem lies in how PISA has come to play such a defining role in determining educational performance and progress. If the educational narrative is dominated by the performance of 15-year-olds in PISA, other important educational goals such as social and emotional development, interpersonal and intrapersonal skills, civics, health and wellbeing, will be held at the margins of the debate. In the end, systems will focus on optimising what they measure.

Looking beyond PISA 2015, there is a substantial opportunity for the OECD to take the lead in navigating the emerging tensions. Two futures are possible. In one future, PISA triggers a global convergence around narrow educational ends, standardising national approaches to curriculum and assessment, and heightening the pressure on policy makers through a highly

publicised three-yearly league table. Yet, an alternative future for PISA is possible. Here the OECD would leverage the strong brand of PISA to lead national governments into a rich discussion about the broad capabilities that matter for learners in the 21st century and the forms of assessment that might provide insight into the success of education systems in developing them. In this future, the OECD would educate policy makers on the limits of indicators, would encourage research that delves into country context and culture in order to understand differential system performance, and would seek to decrease the stakes on policy makers in order to accelerate learning. In this future, PISA would serve to highlight the multi-dimensionality of school systems and support the implementation of coherent and sustainable reform that can impact positively upon the educational and life outcomes of young people and the societies to which they will contribute.

Endnotes

1. Since the release of the first round of PISA results in December 2001, PISA has been a contested survey in the academic literature. At times, the literature has presented a sustained debate between differing perspectives (eg, Prais, 2003, 2004; Adams, 2003; see also special issue of *Assessment in Education: Principles, Policies & Practice*, 2008, Vol 15, p 3). Some scholars have adopted a favourable view of PISA and argued for its benefits in gaining new insights into education system performance (eg, McGaw, 2008b). They reinforce the soundness of the PISA methodology and propose that appropriate steps have been taken to assure validity and reliability of the results. Other scholars have published a series of critiques of PISA focused on both the validity and reliability of the PISA assessment, in addition to the use of the results in ranking countries and drawing out policy conclusions (eg, Prais, 2003, 2004; Goldstein, 2004; Goldstein and Thomas, 2008a and b). It is important to recognise that many of these issues are not unique to PISA, and rather have been raised previously with regards to the IEA studies (see special issue, *Assessment in Education*, 1996, Vol 3, p 2).
2. PIRLS is an international comparative study, conducted every five years, that assesses 4th-grade reading literacy (primary school). TIMSS assesses the mathematics and science achievement of 4th-grade and 8th-grade students (primary and lower-secondary school). The TIMSS survey has been conducted every four years since its first administration in 1995.
3. education.alberta.ca/media/6581166/framework.pdf.

References

- Adams, R J (2003) 'Response to "Cautions on OECD's recent educational survey (PISA)"', *Oxford Review of Education*, 23, p 377–389.
- Bottani, N (1996) 'OECD international education indicators', *International Journal of Educational Research*, 25, 3, p 279–288.
- Breakspear, S (2012) *The Policy Impact of PISA: An Exploration of Normative Effects of International Benchmarking in School System Performance*, OECD Education Working Paper number 71, OECD, Paris.
- Campbell, D T (1976) *Assessing the Impact of Planned Social Change*, Occasional Paper #8, The Public Affairs Center, Dartmouth College, Hanover, New Hampshire, December.
- Commission of the European Communities (2011) *Frequently Asked Questions: Key Report on Progress and Performance in Education and Training in EU Countries*, Press release, European Commission, Brussels.
- Delors, J (1996) *Learning the Treasure Within*, UNESCO, Paris.
- Der Spiegel (no author attributed) (2002) *Dumm gelaufen. Die neue deutsche bildungskatastrophe*, Der Spiegel, Hamburg, 13 May 2002. Accessed 10 August 2014, at www.spiegel.de/spiegel/print/d-22539175.html.
- Egelund, N (2008) 'The value of international comparative studies of achievement – a Danish perspective', *Assessment in Education: Principles, Policy & Practice*, 15, p 245–251.
- Ertl, H (2006) 'Education standards and the changing discourse on education: The reception and consequences of the PISA study in Germany', *Oxford Review of Education*, 32, 5, p 619–634.
- Goldstein, H (2004) 'International comparisons of student attainment: Some issues arising from the PISA study', *Assessment in Education: Principles, Policy & Practice*, 11, p 319–330.
- Goldstein, H and Thomas, S (2008a) 'International comparative studies of achievement – re-examining the issues and impacts', *Assessment in Education: Principles, Policy & Practice*, 15, 3, p 211–213.
- Goldstein, H and Thomas, S (2008b) 'Reflections on the international comparative surveys debate', *Assessment in Education: Principles, Policy & Practice*, 15, 3, p 215–222.
- Henry, M, Lingard, B, Rizvi, F and Taylor, S (2001) *The OECD, Globalisation and Education Policy*, Oxford, Pergamon Press.
- Jakobi, A and Martens, K (2010) 'Expanding and intensifying governance: The OECD in education policy', in K Martens and A Jakobi (2010) *Mechanisms of OECD Governance: International Incentives for National Policy Making*, Oxford University Press, Oxford.
- Kingdon, J (1984) *Agendas, Alternatives, and Public Policy*, HarperCollins Publishers, New York.
- Martens, K and Wolf, K D (2004) *Boomerangs and Trojan Horses – The Unintended Consequences of Internationalising Education Policy through the EU and the OECD*, Springer Publishing, New York.
- Masters, G N (2013) *Reforming Educational Assessment: Imperatives, Principles and Challenges*, Australian Council for Educational Research (ACER), Camberwell, Victoria. Accessed 10 August 2014, at research.acer.edu.au/cgi/viewcontent.cgi?article=1021&context=aer.
- MCEETYA (2008) *Melbourne Declaration on Educational Goals for Young Australians*, Ministerial Council on Education, Employment, Training and Youth Affairs, Melbourne. Accessed 1 August 2014, at www.curriculum.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf.
- McGaw, B (2008a) 'The role of the OECD in international comparative studies of achievement', *Assessment in Education: Principles, Policy & Practice*, 15, p 223–243.
- McGaw, B (2008b) 'Further reflections', *Assessment in Education: Principles, Policy & Practice*, 15, p 279–282.
- Mortimore, P (2008) 'A league table to worry us all', *The Guardian*, 8 January 2008. Accessed 14 November 2014, at www.theguardian.com/profile/petermortimore?page=3.
- OECD (1992) *Education at a Glance*, OECD Publications, Paris.
- OECD (1993) *Education at a Glance*, OECD Publications, Paris.
- OECD (1999) *Measuring Student Knowledge and Skills. A New Framework for Assessment*, OECD Publications, Paris.
- OECD (2000) *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading Mathematics and Science Literacy*, OECD Publications, Paris.
- OECD (2001) *Knowledge and Skills for Life. First Results from PISA*, OECD Publications, Paris.

- OECD (2004) *Learning for Tomorrow's World: First Results from PISA 2003*, OECD Publications, Paris.
- OECD (2010) *PISA 2009 Results: What Students Know and Can Do. Student Performance in Reading, Mathematics and Science. Volume I*, OECD Publications, Paris.
- OECD (2013) *Innovative Learning Environments*, OECD Centre for Educational Research and Innovation (CERI), Paris.
- OECD (2014) *Synergies for Better Learning*, OECD Publications, Paris.
- Phillips, D and Ochs, K (Eds) (2004) *Educational Policy Borrowing: Historical Perspectives*, Symposium Books, Oxford.
- Poovey, M (1998) *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society*, University of Chicago Press, Chicago.
- Prais, S J (2003) 'Cautions on OECD's recent educational survey (PISA)', *Oxford Review of Education*, 29, 2, p 139–163.
- Prais, S J (2004) 'Cautions on OECD's recent educational survey (PISA): Rejoinder to OECD's Response', *Oxford Review of Education*, 30, p 569–573.
- Pring, R (2012) 'Importance of philosophy in the conduct of educational research', *Journal of International and Comparative Education*, 2012, 1, 1, p 23–30.
- Sahlberg, P (2011) *Finnish Lessons: What Can the World Learn From Education Change in Finland*, Teachers College Press, New York.
- Schleicher, A (2007) 'Can competencies assessed by PISA be considered the fundamental school knowledge 15-year-olds should possess?', *Journal of Educational Change*, 8, 349–357.
- Schleicher, A (2009) 'Securing quality and equity in education: Lessons from PISA', *Prospects*, 39, p 251–263.
- Stanley, G (2013) 'Foreword' to G N Masters (2013) *Reforming Educational Assessment: Imperatives, Principles and Challenges*, Australian Council for Educational Research (ACER), Camberwell, Victoria.
- Takayama, K (2008) 'The politics of international league tables: PISA in Japan's achievement crisis debate', *Comparative Education*, 44, p 387–407.
- Takayama, K (2010) 'Politics of externalization in reflexive times: Reinventing Japanese education reform discourses through "Finnish PISA Success"', *Comparative Education Review*, 54, p 51–75.
- The Economist (no author attributed) (2008) 'Education reform. Top of the class. How to learn the right lessons from other countries', *The Economist*, 26 June 2008. Accessed 14 November 2014, at www.economist.com/node/11622383.
- The Economist (no author attributed) (2010) 'Education. An international report card. Shanghai's school students out-perform all others', *The Economist*, 7 December 2010. Accessed 14 November 2014, at www.economist.com/blogs/dailychart/2010/12/education.
- Tucker, M (2011) *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems*, Harvard Education Press, Cambridge, MA.
- Turner, R and Adams, R J (2007) 'The Programme for International Student Assessment: An overview', *Journal of Applied Measurement*, 8, 3, p 237–248.

CSE/IARTV Publications

Recent titles in the Centre for Strategic Education Seminar Series

- No. 240** *How does PISA shape education policy making? Why how we measure learning determines what counts in education*
By Simon Breakspear (November 2014)
- No. 239** *Assessment renaissance*
By Peter Hill and Michael Barber (November 2014)
- No. 238** *An activist or a mature teaching profession? Reflections on progress*
By Judyth Sachs (September 2014)
- No. 237** *Facilitated networks of learning*
By Jean Annan, Brian Annan, Mary Wootton and Rene Burton (September 2014)
- No. 236** *Educators as designers: New systems to scale deep learning*
By Nelson R González (July 2014)
- No. 235** *Australia's national curriculum: A step in the right direction?*
By William Loudon (July 2014)
- No. 234** *A framework for transforming learning in schools: Innovation and the spiral of inquiry*
By Helen Timperley, Linda Kaser and Judy Halbert (April 2014)
- No. 233** *Realigning the governance of schools in Australia: Energising an experimentalist approach*
By Brian Caldwell (April 2014)
- No. 232** *Total Learning*
By Charles Leadbeater (February 2014)
- No. 231** *Can transforming education systems be led?*
By Valerie Hannon (February 2014)
- No. 230** *Reforming an education system: The Tasmanian experience*
By John Ewington and Sally Milbourne with Michael Brakey (November 2013)
- No. 229** *Transforming schooling: Small change is not enough*
By David Loader (November 2013)
- No. 228** *Intercultural Understanding: A key capability in the multicultural world*
Edited by Keith Redman (September 2013)
- No. 227** *Teaching gifted students: A knowing and thinking-based framework for differentiation*
By John Munro (September 2013)
- No. 226** *Paradigm shift: Defining and designing world-class education*
By Yong Zhao (July 2013)
- No. 225** *Gifted students as expert* knowers: A teaching friendly model of gifted knowing and understanding?*
By John Munro (July 2013)
- No. 224** *Exploding the myths of school reform*
By David Hopkins (April 2013)
- No. 223** *Leading the self-transforming school*
By Brian J Caldwell and Jim M Spinks (April 2013)
- No. 222** *The role of professional learning in determining the teaching profession's future*
By Philippa Cordingley (February 2013)
- No. 221** *Building intercultural understanding: What does this mean for students?*
Edited by Keith Redman (February 2013)
- No. 220** *Learning is everyone's business: Educational partnerships between schools, philanthropic, business and third sector organisations*
By Adam Smith (November 2012)
- No. 219** *Australia in the Asian Century: Implications for Australian schooling*
By Kathe Kirby with Dahle Suggett (November 2012)

Other publications

Leading the education debate Volume 3: Selected papers from the CSE's Seminar Series and Occasional Papers, 2007–2010 (2011)
Editors Vic Zbar and Tony Mackay

This third collection from the CSE Seminar Series and Occasional Papers has contributions from a number of significant contemporary and international education writers. It comprises 15 papers on school improvement and reform, and is organised in five parts: The challenge of implementation; Leadership remains key; Improving classroom practice; Disciplined innovation for improved systems and schools; and A system that engages educators, students and the community.

Volume 3 of *Leading the education debate* continues the sequence of volumes of collections of CSE papers. The two earlier volumes by the same editors, *Leading the education debate* (2003) and *Leading the education debate Volume 2* (2007), are also available from CSE.

Women in school leadership: Journeys to success (2010)
Compiled by Jim Watterston

Twelve women reflect on their personal and professional journeys to school leadership, the barriers they have overcome, the successes they have achieved and what they have learned along the way. Their experiences and advice provide inspiration for any teacher who might aspire to school leadership.

About the author

Simon Breakspear works at the intersection of whole-system reform and innovation. As CEO of LearnLabs, he works with system, school and technology leaders to design and implement strategies for deeper learning at scale. Simon holds bachelors degrees in Psychology and Teaching. He did his graduate studies in education policy at the University of Oxford and the University of Cambridge.

About the Paper

The author outlines PISA's inception, design and growth; identifies emerging tensions to be navigated; and explores key implications for system leaders in policy making. He argues that if the educational narrative is dominated by the performance of 15-year-olds in PISA, other important educational goals such as social and emotional development, interpersonal and intrapersonal skills, civics, health and wellbeing, as well as the progression to further work and employment, will be held at the margins of the debate. Looking to the future, he suggests the OECD might educate policy makers on the limits of indicators, and conduct research that delves into country context and culture in order to understand differential system performance.

About the Seminar Series

This series of papers, by leading educators, is based primarily on seminar presentations.

The series is intended to encourage discussion of major issues in education. Views expressed by the authors do not necessarily represent views of the Centre for Strategic Education. Comments on papers are most welcome.

How to order back issues

A complete back catalogue of the CSE/IARTV Seminar and Occasional Paper Series, subscription rates to both of these series and detailed information on other recent publications are available on the Centre for Strategic Education website www.cse.edu.au.

Alternatively contact
Centre for Strategic Education:
phone (+61 3) 9654 1200
fax (+61 3) 9650 5396
email info@cse.edu.au



Leading educational thinking and practice

The constituent bodies of CSE/IARTV are the Association of Heads of Independent Schools of Australia (Vic) and the Victorian Independent Education Union.

ISSN 1838-8558
ISBN 978-1-921823-62-6